The EDRN Knowledge Environment

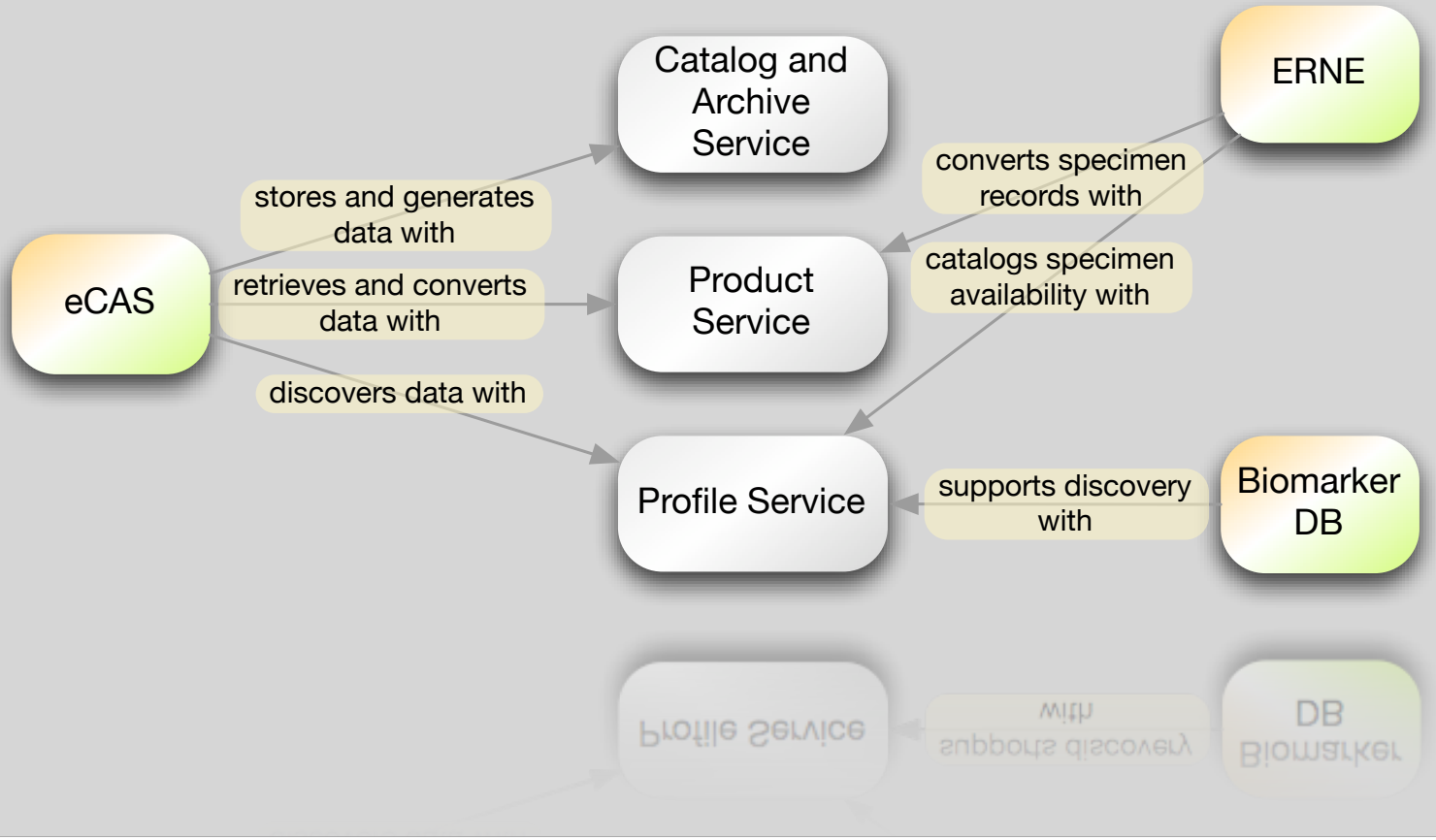
Sean Kelly, Technologist
Object Oriented Data Technology
Jet Propulsion Laboratory
California Institute of Technology
National Aeronautics and Space Administration
Early Detection Research Network
National Cancer Institute
National Institutes of Health

Many Systems

Many commonalities?

Applications for Cancer Research

- Informatics are vital to the progress of cancer research
- EDRN has developed standalone informatics applications
    - Site information, study management tools, protocol data, etc.
- JPL developed an *architecture* for informatics
    - Including applications that are *interoperable*
        - ERNE, eCAS, Biomarker Database, and so forth



Common Components

the architecture

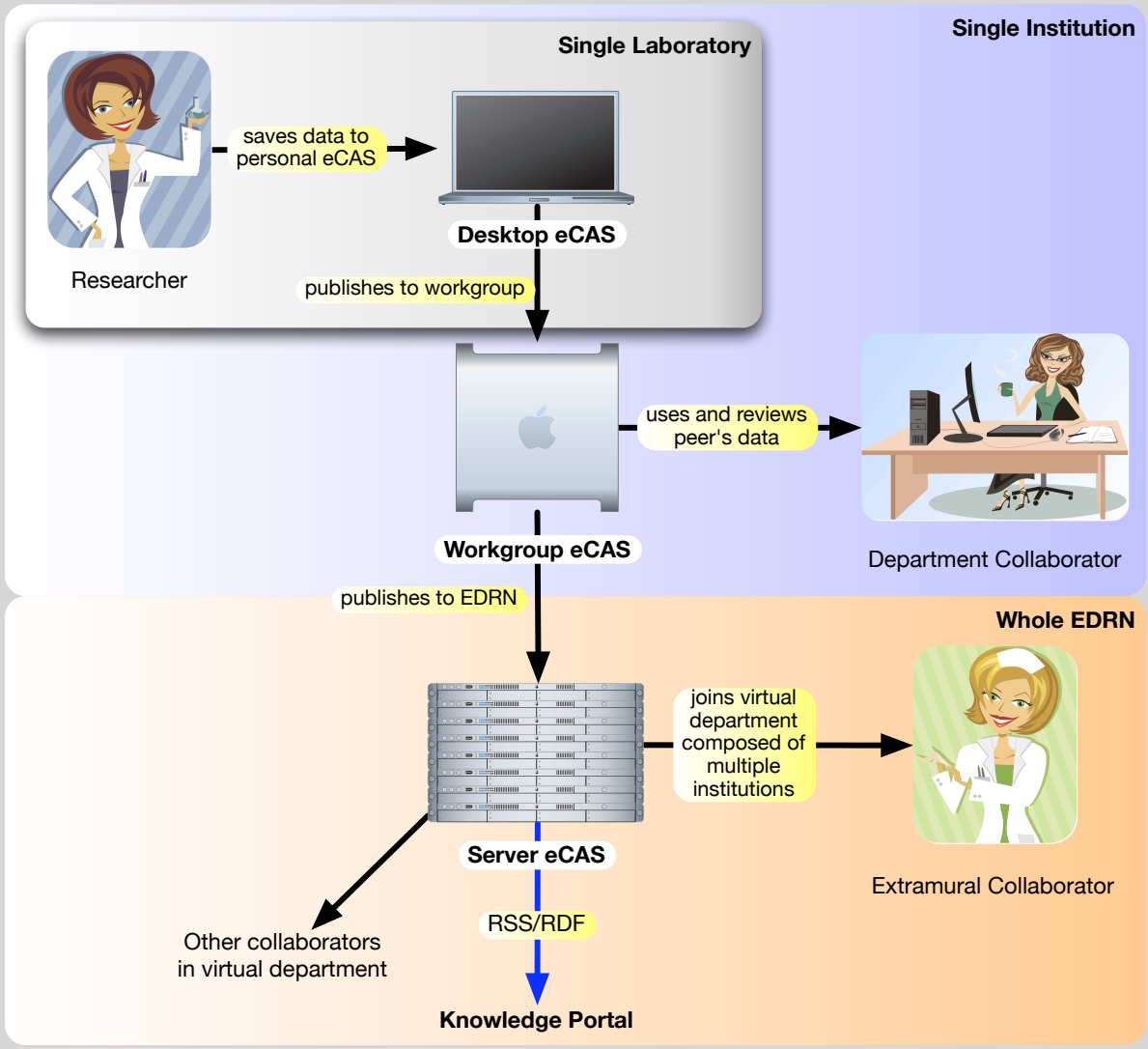
Keys to Interoperability

- Applications must speak a common *metadata language*
  - When you say “37.9”, I need to know what you mean
    - Millimeters of invasiveness, milliliters of blood, body temperature in degrees centigrade, percentage of blast cells in a tissue block, etc.
- When data *share* some set of metadata elements, they may be *correlated*
  - Automatic correlation and correlative searches lead to
    - *Accelerated discovery*

The Semantic Web

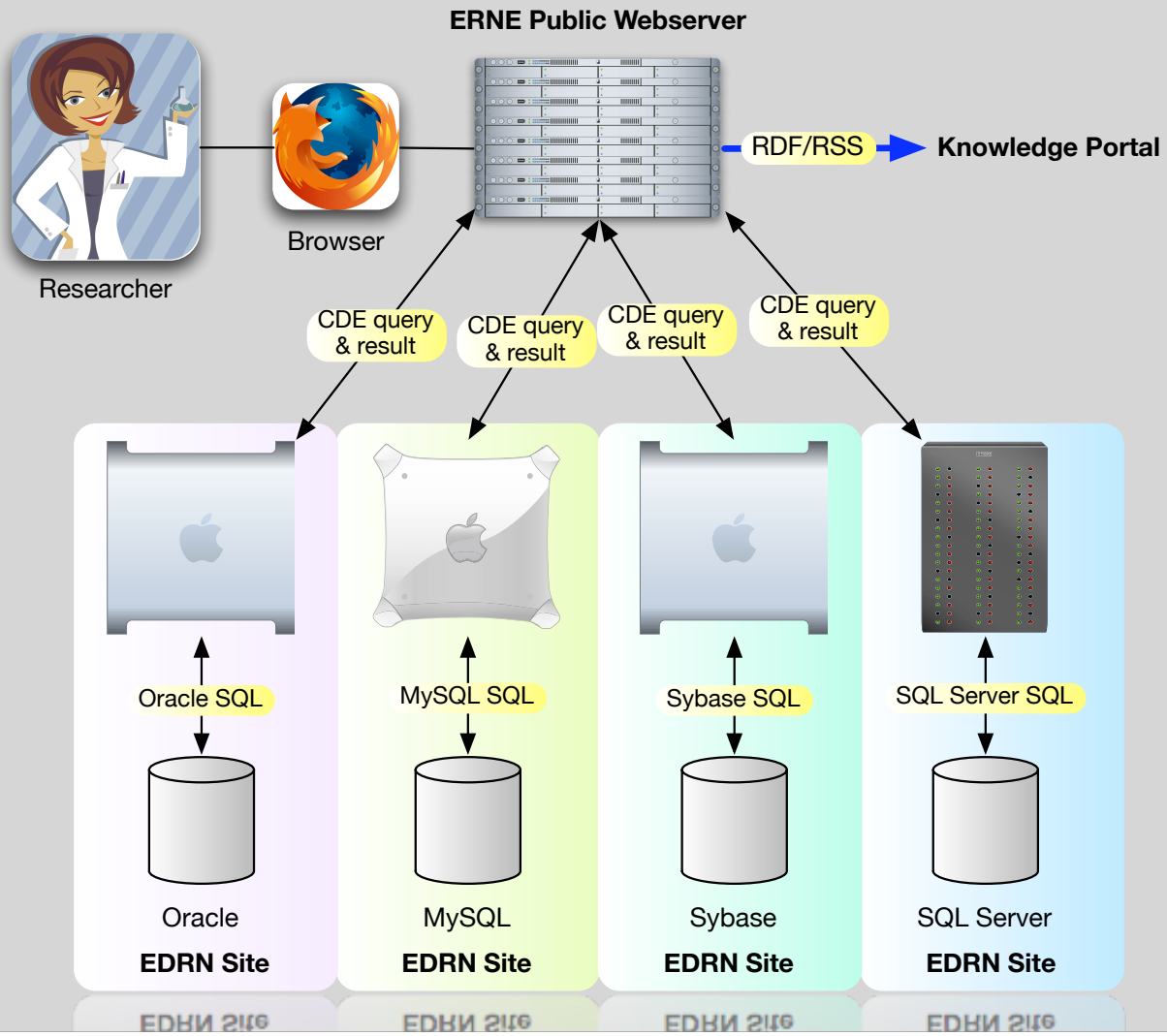
- Attach a URI to any kind of metadata term
    - “cell count” is described at http://some.host/rdfs/schema#cellcount
- Then use the Resource Description Format to describe EDRN entities
    - RDF identifies resources with URIs
    - And describes them with statements
- Specimen at http://some.host/specimens/1192812 has a “cell count” (http://some.host/rdfs/schema#cellcount) of 396

Publish each application's entities in RDF!



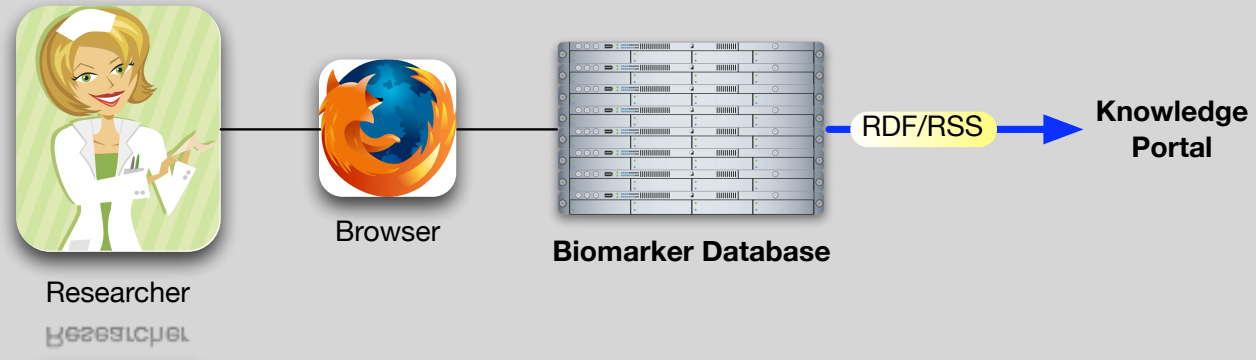
Sending eCAS data

to a single knowledge portal



Sending specimen data

to a single knowledge portal



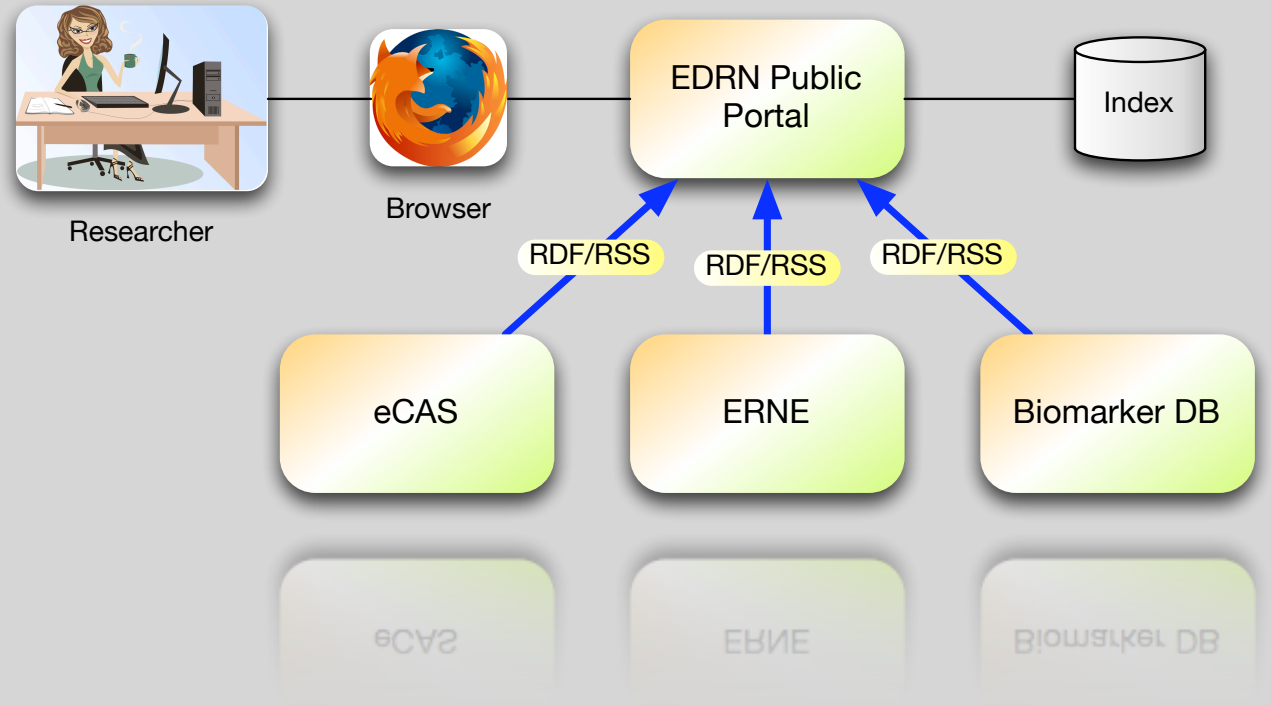
Sending biomarkers

to a single knowledge portal

Ramifications?

- eCAS sites may choose to, or not to, and even *which* data to publish
- ERNE publishes its entire virtual specimen bank
- Biomarker database publishes all biomarkers
- Future applications need only describe entities with RDF
    - And so long as RDFS is available
        - Search forms and help text may be displayed

All available on the EDRN public portal

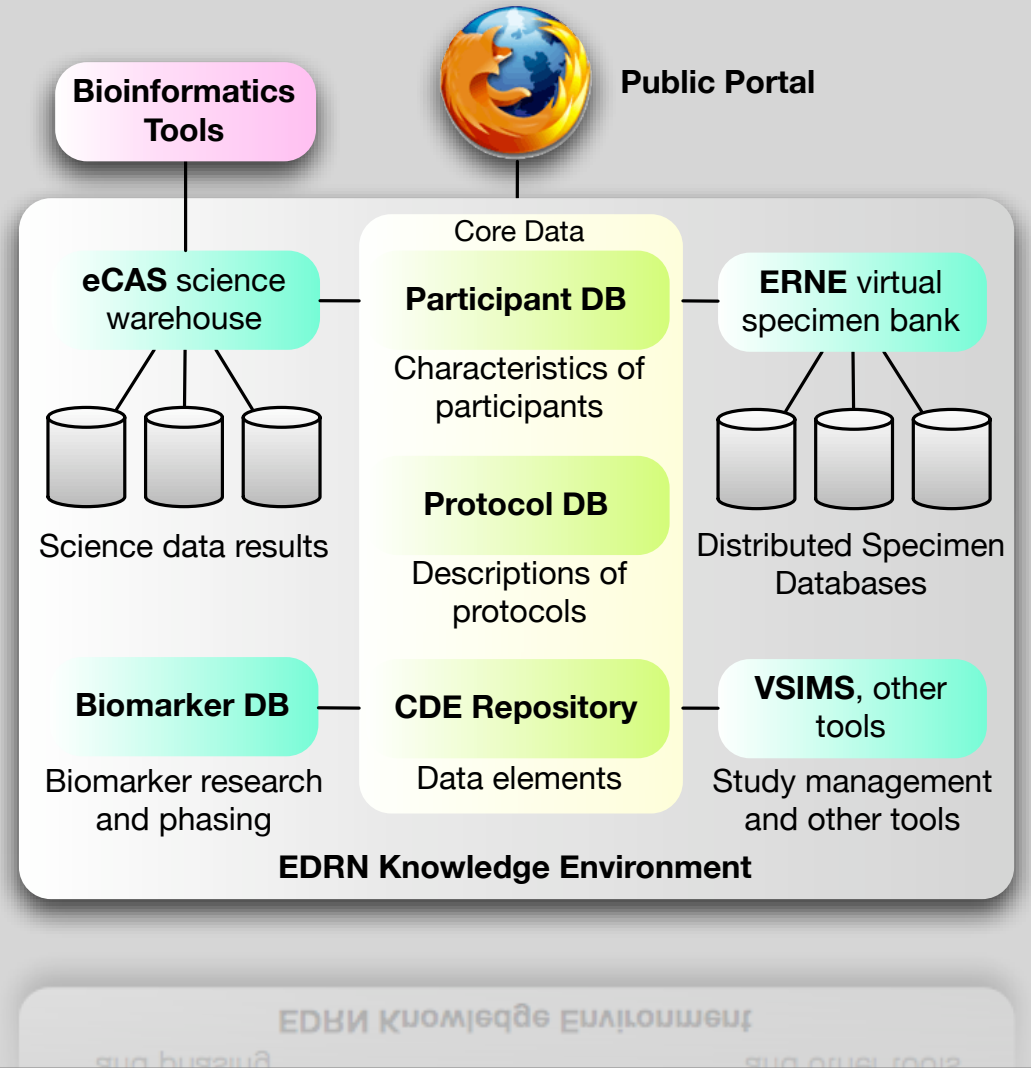


One stop shopping

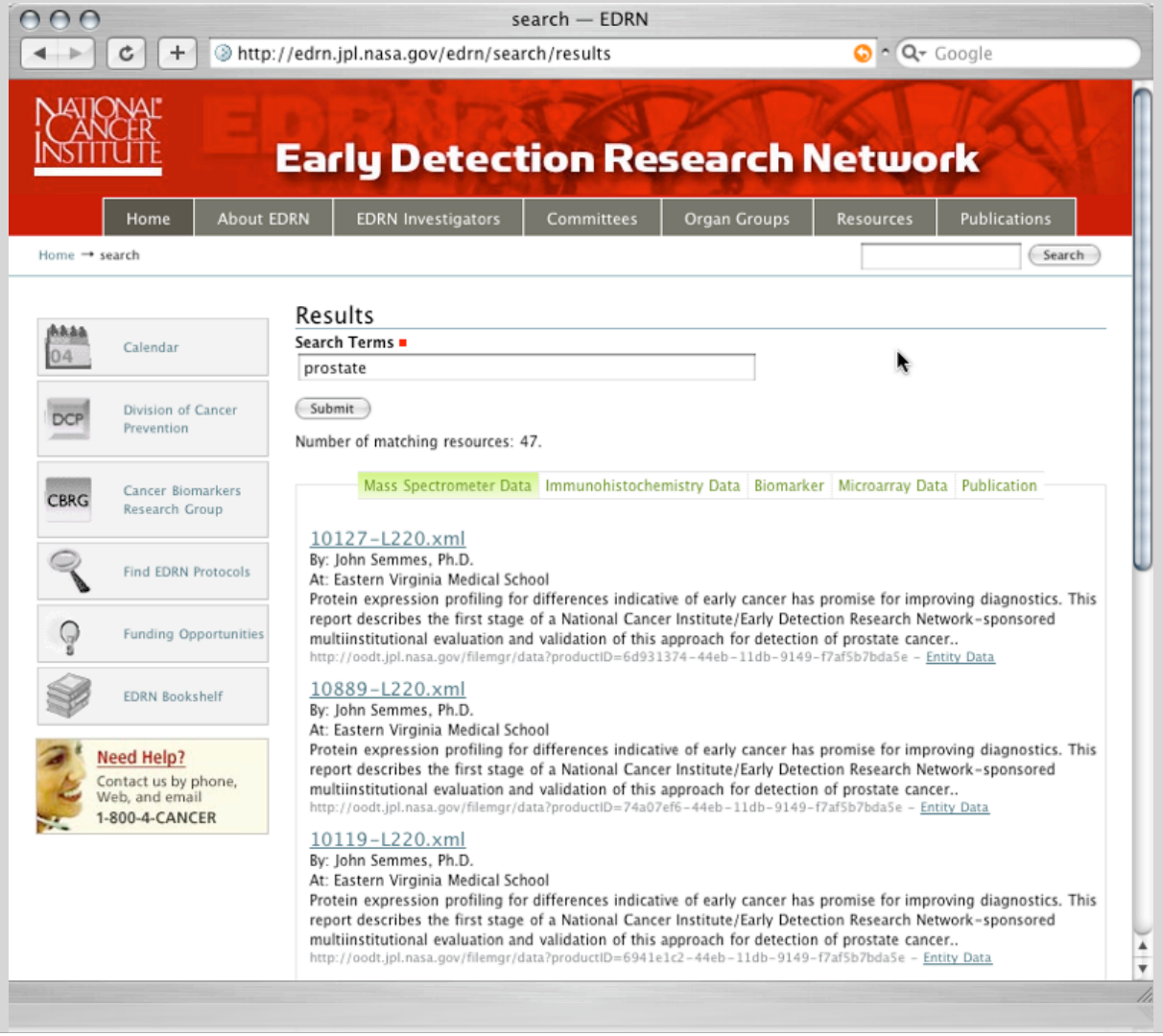
for cancer research data

One-Stop Application Development

- Not just with a browser
- Web services interfaces from portal
  - XML-RPC, REST, and (in future) SOAP
  - Develop new cancer research tools



Prototype | architecture



Prototype

screen shot

Conclusions

- Faster correlations and discoveries
    - Potential for *automated discovery*
- Unification of disparate applications
    - Made possible through *common vocabularies* and RDF
- Way cool idea